

Whole Genome Amplification – review of applications and advances

Hawkins, T. L., Detter, J.C and Richardson, P.
US DOE Joint Genome Institute

Summary of recent advances

The concept of Whole Genome Amplification is something that has arisen in the past few years as modifications to the polymerase chain reaction (PCR) have been adapted to replicate regions of genomes which are of biological interest. The applications here are many – forensics, embryonic disease diagnosis, bio terrorism genome detection, ‘imoralization’ of clinical samples, microbial diversity, and genotyping. The key question is if DNA can be replicated a genome at a time without bias or non random distribution of the target. Several papers published in the last year and currently in preparation may lead to the conclusion that whole genome amplification may indeed be possible and therefore open up a new avenue to molecular biology.

Abbreviations

Polymerase Chain reaction: PCR, Rolling Circle Amplification: RCA, DOP-PCR, Degenerate Oligonucleotide PCR; PEP, Primer Extension Preamplification; CGH, Comparative Genomic Hybridization; PGD; Preimplantation Genetics and Diagnosis

Introduction

The basic premise of molecular biology is that you’re working on DNA. A simple fact, but in many cases perhaps the hardest start point to achieve. The phrase “I lost my DNA” was about as common as “anyone want to get a beer” when I was a graduate student! Today the abundance of kits for DNA manipulation and clean up are ‘almost’ fail safe but the issue still remains that getting good, clean DNA to start with is the most difficult step.

Over the past ten years, methods have been described such as degenerate oligonucleotide PCR (DOP-PCR) (Telenius et al 1992 and Cheung and Nelson 1996) and Primer Extension Preamplification (PEP) (Zhang et al 1992) which have been used as precursors to a variety of genetic tests and assays. Neither DOP-PCR nor PEP claims to replicate the target DNA in its entirety (Cheug & Nelson 1996) or provide a complete coverage of particular loci (Paunio, Reima and Syvanen 1996) but rather provide an amplified source for genotyping or marker identification of some type. It is also the case that the product produced

via these methods is short (<3Kb) and therefore cannot be used in many applications (Telenius, Carter and Bebb et al 1992). These approaches have been proven to be vital in the fields of forensics and genetic disease diagnosis where DNA quantities are limited and hard to find but many tests are required investigating few markers or loci.

One recent example of the widespread use of these approaches is in PGD; preimplantation genetics and diagnosis. This is the basic technique used to infer the status of an embryo from a single cell biopsied 2-3 days after fertilization. Several tests have been described that assay single loci for defects including one that uses multiplex fluorescent PCR to detect trinucleotide amplification in the myotonic dystrophy locus (Dean, Tan and Ao). In a recent paper by Wells and Delhanty whole genome amplification using DOP-PCR has been linked to comparative genomic hybridization (CGH) to assay chromosome copy number changes from single preimplantation embryo cells. The key issue here was the use of whole genome amplification since published methods for CGH required 0.2-1ug of DNA whereas a single cell only contained about 6pg of DNA. In using the combination of these two methods, the authors were able for the first time determine the copy number of every chromosomal region over 10Mb in all of the 12 embryos studied. With the ability to start from a single biopsied cell, this method holds the promise of a very rapid and successful determination of levels of mosaicism, aneuploidy and chromosomal breakage that may explain the low success rates of natural and assisted conception in humans.

A further approach has linked the whole genome amplification methodology of PEP to allow the microsatellite analysis of single tumor cells. Dietmaier et al have described a modified PEP in which improved thermal cycle conditions, cell lysis approaches and the addition of a higher fidelity polymerase have yielded a method which has provided results from single tumor cells allowing microsatellite instability studies, loss of heterozygosity and direct mutation detection via DNA sequencing to be performed. This opens up a new approach in the molecular detection and progression of tumors starting from single biopsied cells.

A new method for whole genome amplification has been introduced in the last year and holds a great deal of promise. The method is based upon the strand displacement amplification approach used in rolling circle amplification (Dean, Nelson, Giesler and Laskin, 2001). Kits are now available for amplification of plasmid templates for DNA sequencing (TempliPhi, Amersham Biosciences) and whole genome amplification systems for research and diagnostics uses are due to follow soon (personal communication John Nelson, Amersham Biosciences & Stephen Kingsmore, Molecular Staging). The method relies upon the high processivity and strand displacement abilities of phi 29 DNA polymerase and the use of random hexamer primers. Initial results have shown that 20-30ug of DNA can be obtained from as few as 1-10 copies of human genomic DNA and that the DNA is large >10Kb (personal communication Stephen Kingsmore, Molecular Staging). Again, this methodology can be used to produce amplified product for

CGH and other genomic analysis. Initial studies using this approach (Dean et al, submitted) has shown that whole genome amplification from human DNA can be performed from a variety of starting materials such as blood or tissue culture cells. To check the integrity of these amplified samples, the authors have used SNP genotyping as well as CGH techniques to show the absence of significant bias or non representation.

One immediate advantage of strand displacement amplification versus PCR based methods may come in phi-29 DNA polymerase error rate. Studies have shown that Phi-29 has an error rate of 1 in $10^6 - 10^7$ (Esteban, Salas and Blanco 1993) in contrast to native taq DNA polymerase at 3 in 10^4 (Eckert and Kunkel 1991) or 1.6×10^{-6} for pfu (Lundberg et al 1991).

The key questions in reference to strand displacement amplification are can this randomly primed strand displacement approach yield a product which is non biased and an accurate replication of the original genome sequence? The answer to these questions may come in the application of microbial genome amplification.

In the field of microbial diversity, the vast majority, perhaps greater than 80%, of microbes are unculturable and many live in communities of multiple organisms. The use of strand displacement amplification could be used to amplify DNA from soil or water samples or from flow sorted samples in the case of microbial communities. The DNA obtained could then be used for DNA sequencing or rapid identification methods. In Detter et al this technique was used to amplify DNA directly from approximately 1000 cells of the microbe *Xylella fastidiosa* yielding over 4ug of high molecular weight DNA once the amplification was completed. Over 30 pcr products throughout the genome were obtained from this amplified DNA. To test for uniformity of coverage of the amplified product, both genomic DNA isolated using standard CsCl based methods and the amplified DNA was made into shotgun libraries for DNA sequencing. Approximately 3000 sequence reads were generated for each totaling 1.2 million high quality bases. These reads were then aligned with the contigs obtained from the whole genome assembly. The amount of coverage for each (39% for the un amplified and 34% for the amplified library) was similar and was in agreement with the expected coverage (38%). These results suggest that the library made from the amplified sample is representative of the genome of the microbe and has not undergone biased amplification.

This experiment shows that DNA amplified by this method from microbial cells without prior isolation of the genomic DNA is a valuable resource for downstream applications including random whole genome shotgun libraries for sequencing. Additionally, indications from sequence information derived from this approach are that the amplified DNA is representative of the original genome, and that the coverage is what is expected from a library sequenced to this depth.

Conclusions

The applications of whole genome amplification, whether using DOP, PEP or strand displacement amplification are gaining ground and opening up new scientific approaches. The ability to take individual cells and generate sufficient, renewable DNA to perform a multitude of genetic tests and assays is the dream of many molecular biologists. For the identification of markers or gross changes in DNA structure, DOP and PEP offer significant advantages and are already in use by large-scale academic and private ventures for genome analysis. Strand displacement amplification offers a different advantage, the ability to regenerate the starting material without bias or non-random distribution of data coverage. In the field of microbial characterization, especially in the current bio terrorism settings, the ability to rapidly amplify genomes from forensic samples and then DNA sequence will allow the vital generation of gene target signatures which can then be used for rapid detection using other technologies. It is also the case that identification of the microbial complement of the human gastrointestinal tract is an important next step in our understanding of the human genome. Technologies like strand displacement amplification could be used to amplify genomes in medical cases where microbial infestation are suggested but not fully supported, such as arthritis.

This work was performed under the auspices of the U.S. Department of Energy, Office of Biological and Environmental Research, by the University of California, under Contracts No. W-7405-Eng-48, No. DE-AC03-76SFOO098, and No. W-7405-ENG-36.

References

Cheung V & Nelson S: **Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA.** *Proc Natl Acad Sci U S A* 1996, **93**: 14676-14679.

Dean F, Nelson J, Giesler T, Lasken R: **Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification.** *Genome Res* 2001, **11**: 1095-1099.

Dean F, Hosono S, Fang L, Wu X, Fawad Faruqi A, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Driscoll M, Song W, Kingsmore S, Egholm M, Lasken R.
Submitted

Dean N, Tan S, Ao A: **The development of preimplantation genetic diagnosis for myotonic dystrophy using multiplex fluorescent polymerase chain reaction and its clinical application.** *Mol Hum Reprod* 2001, **7**: 895-901.

Detter J, Jett J, Arellano A, Ferguson A, Wang M, Nelson J, Lucas S, Chapman J, Lou Y, Rokhsar D, Richardson P, Hawkins T. *Submitted*.

Dietmaier W, Hartmann A, Wallinger S, Heinmoller E, Kerner T, Endl E, Jauch KW, Hofstadter F, Ruschoff J: **Multiple mutation analyses in single tumor cells with improved whole genome amplification**. *Am J Pathol* 1999, **154**: 83-95.

Eckert K, Kunkel, A: **DNA polymerase fidelity and the polymerase chain reaction**. *PCR Methods Appl* 1991, **1**: 17-24.

Esteban J, Salas M. Blanco L: **Fidelity of phi 29 DNA polymerase. Comparison between protein-primed initiation and DNA polymerization**. *J Biol Chem* 1993, **268**: 2719-2726.

Lundberg K, Shoemaker D, Adams M, Short J, Sorge J, and Mathur E: **High-fidelity amplification using a thermostable DNA polymerase isolated from Pyrococcus furiosus**. *Gene* 1991, **108**:1-6.

Telenius H, Carter N, Bebb C, Nordenskjold M, Ponder B & Tunnacliffe A: **Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer**. *Genomics* 1992, **13**: 718-725.

Wells D and Delhanty D: **Comprehensive chromosomal analysis of human preimplantation embryos using whole genome amplification and single cell comparative genomic hybridization**. *Mol Hum Reprod* 2000, **6**: 1055-1062.

Zhang, L, Cui, X, Schmitt, K, Hubert, R, Navidi, W, Arnheim, N: **Whole genome amplification from a single cell: implications for genetic analysis**. *Proc Natl Acad Sci U S A* 1992, **89**: 5847-5851.

Keywords

PCR, Whole Genome Amplification, PEP, Primer Extension Preamplification; CGH, Comparative Genomic Hybridization; PGD; Preimplantation Genetics and Diagnosis.